

# ROAD DETECTION USING DATA FROM MOBILE ROBOT CAMERA

**Jaroslav Peška**

Bachelor Degree Programme (3), FEEC BUT

E-mail: xpeska04@stud.feec.vutbr.cz

Supervised by: Petr Petyovský

E-mail: petyovsky@feec.vutbr.cz

**Abstract:** The paper is focused on developing a road detection algorithm that uses only data from a mobile robot's camera. Key requirements are low latency and relatively low power requirements. Presented algorithm makes use of machine learning, where the neural network is fed not only image data, but also select additional inputs.

**Keywords:** road detection, image processing, neural network, machine learning, obstacle avoidance

## 1 ÚVOD

Jedním z velkých problémů pro automatizaci dopravy je spolehlivá a dostatečně rychlá detekce cesty, po které by se vozidlo mělo pohybovat. Ačkoliv výzkum v této oblasti probíhá již po několik desetiletí, teprve nyní je dostupný dostatečný výpočetní výkon pro praktické nasazení v praxi. Cílem této práce je navrhnout takový algoritmus, který bude primárně zaměřený na podmínky soutěže Robotour, nemělo by však být složité jej adaptovat pro provoz na ostatních komunikacích.

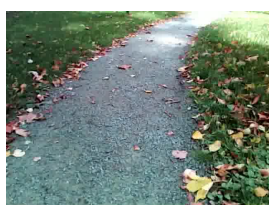
## 2 DETEKCE CESTY

Pro řešení s využitím neuronových sítí jsem se rozhodl, protože v tomto směru existuje ve srovnání s klasickými metodami (tok textury, hledání úběžníku) [2] menší počet publikací.

### 2.1 DATA PRO TESTOVÁNÍ A TRÉNOVÁNÍ ALGORITMU

Pro programatické testování přesnosti algoritmu byla vytvořena sada testovacích dat, nasnímaných ve skutečných podmínkách na prototypu podvozku soutěžního robota vlastní konstrukce.

Pro sběr dat byla využita vícespektrální kamera Intel RealSense R200. Ta obsahuje kromě standardního RGB snímače také stereo infračervené snímače, umožňující případně získat informace o vzdálenosti. [1] Veškerá nasnímaná data (obr. 1) jsou z důvodu snižování paměťových nároků ukládána v rozlišení 640x480 obrazových bodů o frekvenci 30 snímků za sekundu.



(a) viditelné spektrum



(b) levé IR čidlo



(c) pravé IR čidlo



(d) hloubková data

Obrázek 1: Ukázka dat poskytovaných kamerou

Z více než 10 minut zaznamenaných dat byl vybrán vždy poslední snímek každé sekundy, na kterém byla označena cesta. Takto označená data byla rozdělena v poměru 4:1 pro trénování neuronové sítě a pro následné testování.

## 2.2 KONFIGURACE NEURONOVÉ SÍTĚ PRO DETEKCI CESTY

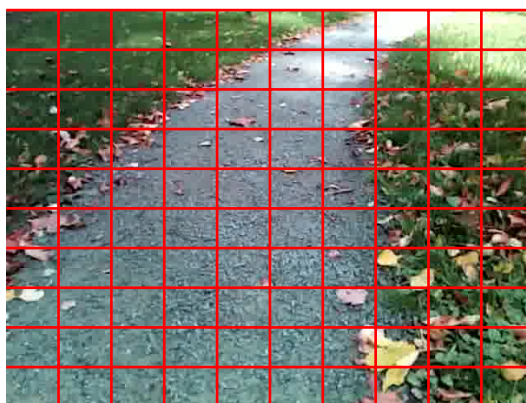
Jedná se o neuronovou síť s jednou skrytou vrstvou o rozměrech  $4/3$  velikosti první vrstvy, v feed-forward konfiguraci (bez zpětných vazeb). Výstupní vrstva obsahuje jediný neuron, udávající zda se v dané sekci cesta nachází či nikoliv. Pro všechny neurony je použita sigmoidální aktivační funkce.

## 2.3 PREPROCESSING DAT PRO DETEKCI CESTY

Nejprve je upravena velikost přijatého obrazu pro rozřezání do částí, viz kapitola 2.4. Následně je odstraněn šum mediánovým filtrem a filtrovaný obraz převeden do vhodnějšího barevného prostoru. Experimentálně bylo zjištěno, že neuronová síť dosahuje nejlepších výsledků při využití barevného prostoru YCrCb. Hloubková data a obraz z infračervených kamer převáděna nejsou, neboť jsou vyjádřena pouze jedním kanálem.

## 2.4 ROZŘEZÁNÍ VSTUPNÍCH DAT DO SEKCI A VÝPOČET

Pro urychlení detekce není obraz zpracováván vcelku, ale je nejprve rozřezán do menších částí. Bez takové optimalizace by nebylo možné z důvodu vysoké složitosti detekovat cestu v téměř reálném čase. Ukázka rozřezání vstupních dat do sekci je uvedena na obrázku 2.



Obrázek 2: Ukázka rozřezání snímku

K obrazovým datům jsou v rámci preprocessingu přidávány přídatné vstupy. V současnosti se jedná zejména o informaci, z jaké části původního obrazu pochází konkrétní sekce. Je tak zajištěno, že neuronová síť bude znát kontext, ze kterého konkrétní část pochází. To umožní například jinou detekci pro sekce z různých částí obrazu. Mezi další přídatné informace mohou patřit mimo jiné data z infračervených kamer. Tyto data jsou velmi přínosné, neboť díky rozdílné míře pohlcení infračerveného světla lze snadno rozlišovat rostlinný porost od cesty, viz obrázek 1.

Takto rozřezaný obraz obohacený o přídatné vstupy je paralelně zpracován. Po zpracování jsou výsledky sestaveny zpět podle jejich pozic v původním obrazu.

### 3 VYHODNOCENÍ ALGORITMU

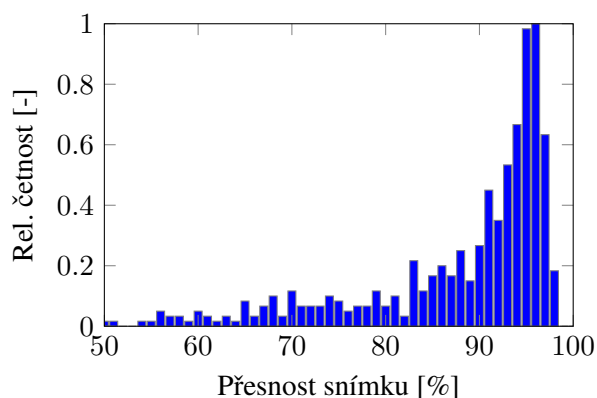
Výsledná přesnost každého snímku je vypočtena jako poměr plochy, kde se výsledek shoduje s etalonem proti celkové ploše obrazu:

$$\text{přesnost} = 1 - \frac{\sum_{x=1}^{\text{šířka}} \sum_{y=1}^{\text{výška}} |\mathbf{etalon}_{y,x} - \mathbf{vypočtený}_{y,x}|}{\text{šířka} \times \text{výška}} \quad [-] \quad (1)$$

kde matice **etalon** je reference a matice **vypočtený** je výstupem algoritmu. Obě matice jsou normalizovány na rozsah hodnot  $\langle 0, 1 \rangle$ .



Obrázek 3: Ukázka výstupu algoritmu



Obrázek 4: Histogram přesnosti výsledků navrženého algoritmu

Navržený algoritmus dosahuje celkové průměrné přesnosti 87,7 % a mediánové přesnosti 92,2 %. Latence celého algoritmu na systému vybaveném procesorem Intel i7-4720HQ (Haswell, 4c/8t) je 45 ms.

### 4 ZÁVĚR

Předvedený algoritmus je plně funkční a při splnění požadavků na rychlost a nenáročnost na výpočetní výkon dosahuje dostatečné přesnosti, což je popsáno v kapitole 3. Ve své bakalářské práci se zaměřím na optimalizaci a využití GPGPU akcelerace, což by mělo přinést několikanásobné zrychlení [3]. Takto získaná rezerva by mohla být následně využita pro snížení počtu částí, na které je obraz rozřezán, což by teoreticky mělo dále zvětšit přesnost.

### REFERENCE

- [1] Intel Corporation, *Intel RealSense Camera R200* [online]. [cit. 2018-03-11]. Dostupné z: <https://software.intel.com/sites/default/files/managed/d7/a9/realsense-camera-r200-product-datasheet.pdf>
- [2] Mikšík, O.: *Dynamic Scene Understanding for Mobile Robot Navigation*. Diplomová práce, Vysoké učení technické v Brně, 2012
- [3] SATO, Kaz, Cliff YOUNG a David Patterson, 2017. An in-depth look at Google's first Tensor Processing Unit (TPU) [online]. May 12, 2017 [cit. 2018-03-13]. Dostupné z: <https://cloud.google.com/blog/big-data/2017/05/an-in-depth-look-at-googles-first-tensor-processing-unit-tpu>